

RESEARCH

Open Access

A class of small deviation theorems for the random fields on an m rooted Cayley tree

Zhiyan Shi^{1*}, Weiguo Yang¹, Lixin Tian¹ and Weicai Peng²

* Correspondence:

shizhiyan1984@126.com

¹Faculty of Science, Jiangsu University, Zhenjiang 212013, China

Full list of author information is available at the end of the article

Abstract

In this paper, we are to establish a class of strong deviation theorems for the random fields relative to m th-order nonhomogeneous Markov chains indexed by an m rooted Cayley tree. As corollaries, we obtain the strong law of large numbers and Shannon-McMillan theorem for m th-order nonhomogeneous Markov chains indexed by that tree.

2000 Mathematics Subject Classification: 60F15; 60J10.

Keywords: strong deviation theorem, m rooted Cayley tree, m th-order nonhomogeneous Markov chain, Shannon-McMillan theorem

1. Introduction

A tree is a graph $G = \{T, E\}$ which is connected and contains no circuits. Given any two vertices $\sigma, t (\sigma \neq t \in T)$, let $\overline{\sigma t}$ be the unique path connecting σ and t . Define the graph distance $d(\sigma, t)$ to be the number of edges contained in the path $\overline{\sigma t}$.

Let $T_{C,N}$ be a Cayley tree. In this tree, the root (denoted by o) has only N neighbors and all other vertices have $N + 1$ neighbors. Let $T_{B,N}$ be a Bethe tree, on which each vertex has $N + 1$ neighboring vertices. Here both $T_{C,N}$ and $T_{B,N}$ are homogeneous tree. In this paper, we mainly consider an m rooted Cayley tree $\overline{T}_{C,N}$ (see Figure 1). It is formed by a Cayley tree $T_{C,N}$ with the root o connecting with another vertex denoted by the root -1 , and then root -1 connecting with another vertex denoted by the root -2 , and continuing to do the same work until the last vertex denoted by the root $-(m - 1)$ is connected. When the context permits, this type of tree is denoted simply by T .

Let $\sigma, t (\sigma, t \neq o, -1, -2, \dots, -(m - 1))$ be vertices of an m rooted Cayley tree T . Write $t \leq \sigma$ if t is on the unique path connecting o to σ , and $|\sigma|$ the number of edges on this path. For any two vertices $\sigma, t (\sigma, t \neq o, -1, -2, \dots, -(m - 1))$ of tree T , denote by $\sigma \wedge t$ the vertex farthest from o satisfying $\sigma \wedge t \leq \sigma$ and $\sigma \wedge t \leq t$.

The set of all vertices with distance n from the root o is called the n -th generation of T , which is denoted by L_n . We say that L_n is the set of all vertices on level n and especially root -1 is on the -1 st level on tree T , root -2 is on the -2 nd level. By analogy, root $-(m - 1)$ is on the $-(m - 1)$ th level. We denote by $T^{(n)}$ the subtree of an m rooted Cayley tree T containing the vertices from level $-(m - 1)$ (the root $-(m - 1)$) to level n . Let $t (t \neq o, -1, -2, \dots, -(m - 1))$ be a vertex of an m rooted Cayley tree T . Predecessor of the vertex t is another vertex, which is nearest from t , on the unique path from root $-(m - 1)$ to t . We denote the predecessor of t by 1_t , the predecessor of 1_t by 2_t and the

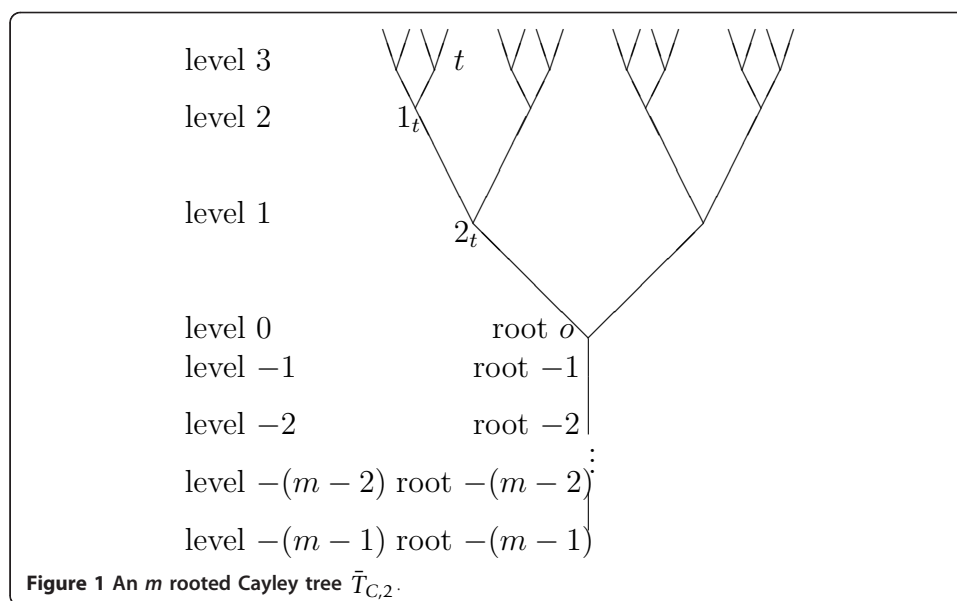


Figure 1 An m rooted Cayley tree $\bar{T}_{C,2}$.

predecessor of $(n-1)_t$ by n_t . We also say that n_t is the n -th predecessor of t . $X^A = \{X_t, t \in A\}$ is a stochastic process indexed by a set A , and denoted by $|A|$ the number of vertices of A , x^A is the realization of X^A .

Let (Ω, \mathcal{F}) be a measure space, $\{X_t, t \in T\}$ be a collection of random variables defined on (Ω, \mathcal{F}) and taking values in $G = \{0, 1, \dots, b-1\}$, where b is a positive integer. Let P be a general probability distribution on (Ω, \mathcal{F}) . We will call P the random field on tree T . Denote the distribution of $\{X_t, t \in T\}$ under the probability measure P by

$$P(x^{T^{(n)}}) = P(X^{T^{(n)}} = x^{T^{(n)}}), \quad x^{T^{(n)}} \in G^{T^{(n)}}. \quad (1)$$

Let

$$f_n(\omega) = -\frac{1}{|T^{(n)}|} \ln P(X^{T^{(n)}}). \quad (2)$$

$f_n(\omega)$ is called entropy density of $X^{T^{(n)}}$.

Let Q be another probability measure on the measurable space (Ω, \mathcal{F}) , and let the distribution of $\{X_t, t \in T\}$ under Q be

$$Q(x^{T^{(n)}}) = Q(X^{T^{(n)}} = x^{T^{(n)}}), \quad x^{T^{(n)}} \in G^{T^{(n)}}. \quad (3)$$

Let

$$h(P|Q) = \limsup_{n \rightarrow \infty} \frac{1}{|T^{(n)}|} \ln \frac{P(X^{T^{(n)}})}{Q(X^{T^{(n)}})}. \quad (4)$$

$h(P|Q)$ is called the sample divergence rate of P relative to Q .

Remark 1 If $P = Q$, $h(P|Q) = 0$ holds. By using the approach of Lemma 1 of Liu and Wang [1], we also can prove that $h(P|Q) \geq 0$, P -a.e.; hence, $h(P|Q)$ can be regarded as a measure of the Markov approximation of the arbitrary random field on T .

Definition 1 (see [2]) Let $G = \{0, 1, \dots, b-1\}$ and $P(y|x_1, x_2, \dots, x_m)$ be a nonnegative functions on G^{m+1} . Let

$$\bar{T}_{C,2}$$

If

$$\sum_{\gamma \in G} P(\gamma | x_1, x_2, \dots, x_m) = 1,$$

then P is called an m -order transition matrix.

Definition 2 (see [2]). Let T be an m rooted Cayley tree, and let $G = \{0, 1, \dots, b-1\}$ be a finite state space, $\{X_t, t \in T\}$ be a collection of G -valued random variables defined on the probability space (Ω, \mathcal{F}, Q) . Let Q be a probability on a measurable space (Ω, \mathcal{F}) .

Let

$$q = (q(x_1, x_2, \dots, x_m)), \quad x_1, x_2, \dots, x_m \in G \quad (5)$$

be a distribution on G^m , and

$$Q_n = (q_n(\gamma | x_1, x_2, \dots, x_m)), \quad x_1, x_2, \dots, x_m, \gamma \in G, n \geq 1 \quad (6)$$

be m -order transition matrices. For any vertex $t \in L_m$, $n \geq 1$, if

$$\begin{aligned} Q(X_t = \gamma | X_{1_t} = x_1, X_{2_t} = x_2, \dots, X_{m_t} = x_m \text{ and } X_\sigma \text{ for } \sigma \wedge t \leq 1_t) \\ = Q(X_t = \gamma | X_{1_t} = x_1, X_{2_t} = x_2, \dots, X_{m_t} = x_m) \\ = q_n(\gamma | x_1, x_2, \dots, x_m), \forall x_1, x_2, \dots, x_m, \gamma \in G \end{aligned} \quad (7)$$

and

$$\begin{aligned} Q(X_{-(m-1)} = x_1, \dots, X_{-1} = x_{m-1}, X_0 = x_m) \\ = q(x_1, \dots, x_{m-1}, x_m), \quad x_1, \dots, x_m \in G, \end{aligned} \quad (8)$$

then $\{X_t, t \in T\}$ is called a G -valued m th-order nonhomogeneous Markov chain indexed by an m rooted Cayley tree with the initial m dimensional distribution (5) and m -order transition matrices (6) under the probability measure Q , or called a T -indexed m th-order nonhomogeneous Markov chain under the probability measure Q .

We denote

$$o_m = \{0, -1, -2, \dots, -(m-1)\}, o'_m = \{-1, -2, \dots, -(m-1)\},$$

$$X_1^n(t) = \{X_{n_t}, \dots, X_{2_t}, X_{1_t}\}, X_0^n(t) = \{X_{n_t}, \dots, X_{2_t}, X_{1_t}, X_t\},$$

and denote by $x_1^n(t)$ and $x_0^n(t)$ the realizations $X_1^n(t)$ and $X_0^n(t)$, respectively.

Let $\{X_t, t \in T\}$ be an m th-order nonhomogeneous Markov chains indexed by an m rooted Cayley tree T under the probability measure Q defined on above. It is easy to see that

$$Q(x^{T(n)}) = Q(x^{T(n)} = x^{T(n)}) = q(x_{-(m-1)}, \dots, x_0) \prod_{k=1}^n \prod_{t \in L_k} q_k(x_t | x_1^m(t)). \quad (9)$$

In the following, we always assume that $P(x^{T(n)})$, $Q(x^{T(n)})$, $q(x_1, \dots, x_m)$, and $\{q_n(\gamma | x_1, \dots, x_m), n \geq 1\}$ are all positive.

There have been some works on limit theorems for tree-indexed stochastic process. Benjamini and Peres [3] have given the notion of the tree-indexed Markov chains and studied the recurrence and ray-recurrence for them. Berger and Ye [4] have studied the existence of entropy rate for some stationary random fields on a homogeneous tree. Pemantle [5] proved a mixing property and a weak law of large numbers for a PPG-invariant and ergodic random field on a homogeneous tree. Ye and Berger [6,7], by using Pemantle's result and a combinatorial approach, have studied the Shannon-McMillan theorem with convergence in probability for a PPG-invariant and ergodic random field on a homogeneous tree. Yang and Liu [8] have studied a strong law of large numbers for the frequency of occurrence of states for Markov chains field on a Bethe tree (a particular case of tree-indexed Markov chains field and PPG-invariant random field). Yang [9] has studied the strong law of large numbers for frequency of occurrence of state and Shannon-McMillan theorem for homogeneous Markov chains indexed by a homogeneous tree. Yang and Ye [10] have studied the strong law of large numbers and Shannon-McMillan theorem for nonhomogeneous Markov chains indexed by a homogeneous tree. Huang and Yang [11] have studied the strong law of large numbers and Shannon-McMillan theorem for Markov chains indexed by an infinite tree with uniformly bounded degree. Recently, Shi and Yang [12] have also studied some limit properties of random transition probability for second-order nonhomogeneous Markov chains indexed by a tree. Peng et al. [13] have studied a class of strong deviation theorems for the random fields relative to homogeneous Markov chains indexed by a homogeneous tree. Shi and Yang [2] have studied the strong law of large numbers and Shannon-McMillan for the m th-order nonhomogeneous Markov chains indexed by an m rooted Cayley tree. Yang [14] has also studied a class of small deviation theorems for the sequences of N -valued random variables with respect to m th-order nonhomogeneous Markov chains.

In this paper, our main purpose is to extend Yang's [14] result to an m rooted Cayley tree. By introducing the sample divergence rate of any probability measure with respect to m th-order nonhomogeneous Markov measure on an m rooted Cayley tree, we establish a class of strong deviation theorems for the arbitrary random fields indexed by that tree with respect to m th-order nonhomogeneous Markov chains indexed by that tree. As corollaries, we obtain the strong law of large numbers and Shannon-McMillan theorem for m th-order nonhomogeneous Markov chains indexed by that tree.

2. Main Results

Before giving the main results, we begin with a lemma.

Lemma 1 Let T be an m rooted Cayley tree, $G = \{0, 1, \dots, b-1\}$ be the finite state space. Let $\{X_t, t \in T\}$ be a collection of G -valued random variables defined on the measurable space (Ω, \mathcal{F}) . Let P and Q be two probability measures on the measurable space (Ω, \mathcal{F}) , and let $\{X_t, t \in T\}$ be an m th-order nonhomogeneous Markov chains indexed by tree T under probability measure Q . Let $\{g_n(y_1, \dots, y_{m+1}), n \geq 1\}$ be a sequence of functions defined on G^{m+1} . Let $\mathcal{F}_n = \sigma(X^{T(n)})(n \geq 1)$. Set

$$F_n(\omega) = \sum_{k=1}^n \sum_{t \in L_k} g_k(X_0^m(t)) \quad (10)$$

and

$$t_n(\lambda, \omega) = \frac{e^{\lambda F_n(\omega)}}{\prod_{k=1}^n \prod_{t \in L_k} E_Q [e^{\lambda g_k(X_0^m(t)) | X_1^m(t)}]} \cdot \frac{q(X_{-(m-1)}, \dots, X_0) \prod_{k=1}^n \prod_{t \in L_k} q_k(X_t | X_1^m(t))}{P(X^{T(n)})}, \quad (11)$$

where E_Q denote the expectation under probability measure Q . Then $\{t_n(\lambda, \omega), \mathcal{F}_n, n \geq 1\}$ is a nonnegative martingale under probability measure P .

Proof The proof is similar to Lemma 3 of Peng et al. [12], so the proof is omitted.

Theorem 1 Let T be an m rooted Cayley tree, $\{X_t, t \in T\}$ be a collection of random variables taking values in $G = \{0, 1, \dots, b-1\}$ defined on the measurable space (Ω, \mathcal{F}) . Let P and Q be two probability measures on the measurable space (Ω, \mathcal{F}) , such that $\{X_t, t \in T\}$ is an m th-order nonhomogeneous Markov chain indexed by T under Q . Let $h(P | Q)$ be defined by (4), $\{g_n(y_1, \dots, y_{m+1}), n \geq 1\}$ be a sequence of functions defined on G^{m+1} . Let $c \geq 0$ be a constant. Set

$$D(c) = \{\omega : h(P|Q) \leq c\}. \quad (12)$$

Assume that there exists $\alpha > 0$, such that $\forall i^m \in G^m$,

$$b_\alpha(i^m) = \limsup_{n \rightarrow \infty} \frac{1}{|T(n)|} \sum_{k=1}^n \sum_{t \in L_k} E_Q [e^{a|g_k(X_0^m(t))|} |X_1^m(t) = i^m] \leq \tau. \quad (13)$$

Let

$$A_t = \frac{2\tau}{e^2(t - \alpha)^2}, \quad (14)$$

where $0 < t < a$. Thus, when $0 \leq c \leq t^2 A_t$, we have

$$\limsup_{n \rightarrow \infty} \frac{1}{|T(n)|} \left| \sum_{k=1}^n \sum_{t \in L_k} \{g_k(X_0^m(t)) - E_Q[g_k(X_0^m(t)) | X_1^m(t)]\} \right| \limsup \leq 2\sqrt{cA_t}, \quad P-a.e., \omega \in D(c). \quad (15)$$

In particular,

$$\lim_{n \rightarrow \infty} \frac{1}{|T(n)|} \sum_{k=1}^n \sum_{t \in L_k} \{g_k(X_0^m(t)) - E_Q[g_k(X_0^m(t)) | X_1^m(t)]\} = 0, \quad P-a.e., \omega \in D(0). \quad (16)$$

Proof Let $t_n(\lambda, \omega)$ be defined by (11). By Lemma 1, $\{t_n(\lambda, \omega), \mathcal{F}_n, n \geq 1\}$ is a non-negative martingale under probability measure P . By Doob's martingale convergence theorem, we have

$$\lim_{n \rightarrow \infty} t_n(\lambda, \omega) = t(\lambda, \omega) < \infty, \quad P-a.e.$$

Hence,

$$\limsup_{n \rightarrow \infty} \frac{1}{|T(n)|} \ln t_n(\lambda, \omega) \leq 0, \quad P-a.e.. \quad (17)$$

We have by (9), (10), (11) and (17)

$$\limsup_{n \rightarrow \infty} \frac{1}{|T(n)|} \left[\sum_{k=1}^n \sum_{t \in L_k} \left\{ \lambda g_k(X_0^m(t)) - \ln E_Q [e^{\lambda g_k(X_0^m(t)) | X_1^m(t)}] \right\} - \ln \frac{P(X^{T(n)})}{Q(X^{T(n)})} \right] \leq 0, \quad P-a.e. \quad (18)$$

By (4),(12) and (18)

$$\limsup_{n \rightarrow \infty} \frac{1}{|T(n)|} \sum_{k=1}^n \sum_{t \in L_k} \left\{ \lambda g_k(X_0^m(t)) - \ln E_Q \left[e^{\lambda g_k(X_0^m(t))} |X_1^m(t)| \right] \right\} \leq c, \quad P-a.e., \omega \in D(c). \quad (19)$$

This implies that

$$\begin{aligned} & \limsup_{n \rightarrow \infty} \frac{\lambda}{|T(n)|} \sum_{k=1}^n \sum_{t \in L_k} \{g_k(X_0^m(t)) - E_Q[g_k(X_0^m(t))|X_1^m(t)]\} \\ & \leq \limsup_{n \rightarrow \infty} \frac{1}{|T(n)|} \sum_{k=1}^n \sum_{t \in L_k} \left\{ \ln E_Q \left[e^{\lambda g_k(X_0^m(t))} |X_1^m(t)| \right] - E_Q[\lambda g_k(X_0^m(t))|X_1^m(t)] \right\} + c, \quad P-a.e., \omega \in D(c) \end{aligned} \quad (20)$$

Let $|\lambda| < t$. By inequalities $\ln x \leq x - 1$ ($x > 0$) and $e^x - 1 - x \leq \frac{x^2}{2}e^{|x|}$, and noticing that

$$\max\{x^2 e^{-hx}, x \geq 0\} = 4e^{-2}/h^2 \quad (h > 0). \quad (21)$$

We have

$$\begin{aligned} & \limsup_{n \rightarrow \infty} \frac{1}{|T(n)|} \sum_{k=1}^n \sum_{t \in L_k} \left\{ \ln E_Q \left[e^{\lambda g_k(X_0^m(t))} |X_1^m(t)| \right] - E_Q[\lambda g_k(X_0^m(t))|X_1^m(t)] \right\} \\ & \leq \limsup_{n \rightarrow \infty} \frac{1}{|T(n)|} \sum_{k=1}^n \sum_{t \in L_k} \left\{ E_Q \left[e^{\lambda g_k(X_0^m(t))} |X_1^m(t)| \right] - 1 - E_Q[\lambda g_k(X_0^m(t))|X_1^m(t)] \right\} \\ & \leq \frac{\lambda^2}{2} \limsup_{n \rightarrow \infty} \frac{1}{|T(n)|} \sum_{k=1}^n \sum_{t \in L_k} E_Q \left[g_k^2(X_0^m(t)) e^{|\lambda| |g_k(X_0^m(t))|} |X_1^m(t)| \right] \\ & = \frac{\lambda^2}{2} \limsup_{n \rightarrow \infty} \frac{1}{|T(n)|} \sum_{k=1}^n \sum_{t \in L_k} E_Q \left[e^{\alpha |g_k(X_0^m(t))|} g_k^2(X_0^m(t)) e^{(|\lambda| - \alpha) |g_k(X_0^m(t))|} |X_1^m(t)| \right] \\ & \leq \frac{\lambda^2}{2} \limsup_{n \rightarrow \infty} \frac{1}{|T(n)|} \sum_{k=1}^n \sum_{t \in L_k} E_Q \left[e^{\alpha |g_k(X_0^m(t))|} 4e^{-2}/(|\lambda| - \alpha)^2 |X_1^m(t)| \right] \\ & \leq 2\lambda^2 \tau / e^2 (t - \alpha)^2. \end{aligned} \quad (22)$$

By (20) and (22), we have

$$\begin{aligned} & \limsup_{n \rightarrow \infty} \frac{\lambda}{|T(n)|} \sum_{k=1}^n \sum_{t \in L_k} \{g_k(X_0^m(t)) - E_Q[g_k(X_0^m(t))|X_1^m(t)]\} \\ & \leq \lambda^2 A_t + c, \quad P-a.e., \omega \in D(c). \end{aligned} \quad (23)$$

When $0 < \lambda < t < \alpha$, we have by (23)

$$\begin{aligned} & \limsup_{n \rightarrow \infty} \frac{1}{|T(n)|} \sum_{k=1}^n \sum_{t \in L_k} \{g_k(X_0^m(t)) - E_Q[g_k(X_0^m(t))|X_1^m(t)]\} \\ & \leq \lambda A_t + c/\lambda, \quad P-a.e., \omega \in D(c). \end{aligned} \quad (24)$$

It is easy to see that when $0 < c < t^2 A_t$, the function $f(\lambda) = \lambda A_t + c/\lambda$ attains, at $\lambda = \sqrt{c/A_t}$, its smallest value $f(\sqrt{c/A_t}) = 2\sqrt{cA_t}$. Letting $\lambda = \sqrt{c/A_t}$ in (24), we have

$$\limsup_{n \rightarrow \infty} \frac{1}{|T(n)|} \sum_{k=1}^n \sum_{t \in L_k} \{g_k(X_0^m(t)) - E_Q[g_k(X_0^m(t))|X_1^m(t)]\} \leq 2\sqrt{cA_t}, \quad P-a.e., \omega \in D(c). \quad (25)$$

When $c = 0$, we have by (24)

$$\limsup_{n \rightarrow \infty} \frac{1}{|T(n)|} \sum_{k=1}^n \sum_{t \in L_k} \{g_k(X_0^m(t)) - E_Q[g_k(X_0^m(t))|X_1^m(t)]\} \leq \lambda A_t, \quad P-a.e., \omega \in D(0). \quad (26)$$

Letting $\lambda \rightarrow 0^+$ in (26), we obtain

$$\limsup_{n \rightarrow \infty} \frac{1}{|T(n)|} \sum_{k=1}^n \sum_{t \in L_k} \{g_k(X_0^m(t)) - E_Q[g_k(X_0^m(t))|X_1^m(t)]\} \leq 0, \quad P-a.e., \omega \in D(0). \quad (27)$$

Hence, (25) also holds for $c = 0$. When $-\alpha < -t < \lambda < 0$, by virtue of (23) it can be shown in a similar way that

$$\liminf_{n \rightarrow \infty} \frac{1}{|T(n)|} \sum_{k=1}^n \sum_{t \in L_k} \{g_k(X_0^m(t)) - E_Q[g_k(X_0^m(t))|X_1^m(t)]\} \geq -2\sqrt{cA_t}, \quad P-a.e., \omega \in D(c). \quad (28)$$

Equation 15 follows from (25) and (28), Equation 15 implies (16) immediately. This completes the proof of the theorem. \square

Theorem 2 Let

$$H_t = 2b/e^2(t-1)^2, \quad 0 < t < 1. \quad (29)$$

Let $f_n(\omega)$ be defined by (2). Under the conditions of Theorem 1, when $0 \leq c \leq t^2 H_t$, we have

$$\begin{aligned} & \limsup_{n \rightarrow \infty} \{f_n(\omega) - \frac{1}{|T(n)|} \sum_{k=1}^n \sum_{t \in L_k} H[q_k(0|X_1^m(t)), \dots, q_k(b-1|X_1^m(t))]\} \\ & \leq 2\sqrt{cH_t}, \quad P-a.e., \omega \in D(c), \end{aligned} \quad (30)$$

$$\begin{aligned} & \liminf_{n \rightarrow \infty} \{f_n(\omega) - \frac{1}{|T(n)|} \sum_{k=1}^n \sum_{t \in L_k} H[q_k(0|X_1^m(t)), \dots, q_k(b-1|X_1^m(t))]\} \\ & \geq -2\sqrt{cH_t} - c, \quad P-a.e., \omega \in D(c), \end{aligned} \quad (31)$$

where $H(p_0, \dots, p_{b-1})$ denote the entropy of distribution (p_0, \dots, p_{b-1}) , i.e.,

$$H(p_0, \dots, p_{b-1}) = - \sum_{i=0}^{b-1} p_i \ln p_i.$$

Proof In Theorem 1, let $g_k(y_1, \dots, y_{m+1}) = -\ln q_k(y_{m+1} | y_1, \dots, y_m)$ and $\alpha = 1$, we have

$$\begin{aligned} & E_Q \left[e^{g_k(X_0^m(t))} | X_1^m(t) = i^m \right] \\ & = \sum_{j \in G} e^{|\ln q_k(j|i^m)|} q_k(j|i^m) \\ & = \sum_{j \in G} q_k(j|i^m) / q_k(j|i^m) \\ & = b. \end{aligned} \quad (32)$$

Hence, $\forall i^m \in G^m$,

$$b_1(i^m) = \limsup_{n \rightarrow \infty} \frac{1}{|T(n)|} \sum_{k=1}^n \sum_{t \in L_k} E_Q \left[e^{g_k(X_0^m(t))} |X_1^m(t) = i^m \right] \leq b. \quad (33)$$

Noticing that

$$\begin{aligned} & E_Q[-\ln q_k(X_t | X_1^m(t)) | X_1^m(t)] \\ &= - \sum_{j \in G} q_k(j | X_1^m(t)) \ln q_k(j | X_1^m(t)) \\ &= H[q_k(0 | X_1^m(t)), \dots, q_k(b-1 | X_1^m(t))] . \end{aligned} \quad (34)$$

When $0 \leq c \leq t^2 H_t$, we have by (34), (29) and (15)

$$\begin{aligned} & \limsup_{n \rightarrow \infty} \left\{ \frac{1}{|T(n)|} \sum_{k=1}^n \sum_{t \in L_k} (-\ln q_k(X_t | X_1^m(t))) - \frac{1}{|T(n)|} \sum_{k=1}^n \sum_{t \in L_k} H[q_k(0 | X_1^m(t)), \dots, q_k(b-1 | X_1^m(t))] \right\} \\ & \leq 2\sqrt{cH_t}, \quad P - a.e., \omega \in D(c). \end{aligned} \quad (35)$$

$$\begin{aligned} & \liminf_{n \rightarrow \infty} \left\{ \frac{1}{|T(n)|} \sum_{k=1}^n \sum_{t \in L_k} (-\ln q_k(X_t | X_1^m(t))) - \frac{1}{|T(n)|} \sum_{k=1}^n \sum_{t \in L_k} H[q_k(0 | X_1^m(t)), \dots, q_k(b-1 | X_1^m(t))] \right\} \\ & \geq -2\sqrt{cH_t}, \quad P - a.e., \omega \in D(c). \end{aligned} \quad (36)$$

By (35), (9) and $h(P|Q) \geq 0$,

$$\begin{aligned} & \limsup_{n \rightarrow \infty} \left\{ f_n(\omega) - \frac{1}{|T(n)|} \sum_{k=1}^n \sum_{t \in L_k} H[q_k(0 | X_1^m(t)), \dots, q_k(b-1 | X_1^m(t))] \right. \\ & \leq \limsup_{n \rightarrow \infty} \left\{ -\frac{1}{|T(n)|} \ln P(X^{T(n)}) - \frac{1}{|T(n)|} \sum_{k=1}^n \sum_{t \in L_k} (-\ln q_k(X_t | X_1^m(t))) \right\} \\ & + \limsup_{n \rightarrow \infty} \left\{ \frac{1}{|T(n)|} \sum_{k=1}^n \sum_{t \in L_k} (-\ln q_k(X_t | X_1^m(t))) \right. \\ & \left. - \frac{1}{|T(n)|} \sum_{k=1}^n \sum_{t \in L_k} H[q_k(0 | X_1^m(t)), \dots, q_k(b-1 | X_1^m(t))] \right\} \\ & \leq 2\sqrt{cH_t}, \quad P - a.e., \omega \in D(c). \end{aligned} \quad (37)$$

By (36), (9) and (12), we have

$$\begin{aligned} & \liminf_{n \rightarrow \infty} \left\{ f_n(\omega) - \frac{1}{|T(n)|} \sum_{k=1}^n \sum_{t \in L_k} H[q_k(0 | X_1^m(t)), \dots, q_k(b-1 | X_1^m(t))] \right. \\ & \geq \liminf_{n \rightarrow \infty} \left\{ -\frac{1}{|T(n)|} \ln P(X^{T(n)}) - \frac{1}{|T(n)|} \sum_{k=1}^n \sum_{t \in L_k} (-\ln q_k(X_t | X_1^m(t))) \right\} \\ & + \liminf_{n \rightarrow \infty} \left\{ \frac{1}{|T(n)|} \sum_{k=1}^n \sum_{t \in L_k} (-\ln q_k(X_t | X_1^m(t))) \right. \\ & \left. - \frac{1}{|T(n)|} \sum_{k=1}^n \sum_{t \in L_k} H[q_k(0 | X_1^m(t)), \dots, q_k(b-1 | X_1^m(t))] \right\} \\ & \geq -h(P|Q) - 2\sqrt{cH_t} \\ & \geq -2\sqrt{cH_t} - c, \quad P - a.e., \omega \in D(c). \end{aligned} \quad (38)$$

This completes the proof of this theorem. \square

Corollary 1 Under the conditions of Theorem 2, we have

$$\lim_{n \rightarrow \infty} \left\{ f_n(\omega) - \frac{1}{|T(n)|} \sum_{k=1}^n \sum_{t \in L_k} H[q_k(0|X_1^m(t)), \dots, q_k(b-1|X_1^m(t))] \right\} = 0, \quad P-a.e., \omega \in D(0). \quad (39)$$

If $P << Q$, then

$$\lim_{n \rightarrow \infty} \left\{ f_n(\omega) - \frac{1}{|T(n)|} \sum_{k=1}^n \sum_{t \in L_k} H[q_k(0|X_1^m(t)), \dots, q_k(b-1|X_1^m(t))] \right\} = 0, \quad P-a.e. \quad (40)$$

In particular, if $P = Q$,

$$\lim_{n \rightarrow \infty} \left\{ f_n(\omega) - \frac{1}{|T(n)|} \sum_{k=1}^n \sum_{t \in L_k} H[q_k(0|X_1^m(t)), \dots, q_k(b-1|X_1^m(t))] \right\} = 0, \quad Q-a.e. \quad (41)$$

Proof Letting $c = 0$ in (30) and (31), Equation 39 follows. If $P << Q$, then $h(P | Q) = 0$, $P-a.e.$, (cf. see [15], P.121), i.e., $P(D(0)) = 1$. Hence, Equation 40 follows from (39). In particular, if $P = Q$, then $h(P | Q) \equiv 0$. Hence, (41) follows from (40). \square

Theorem 3 Under the conditions of Theorem 1, if $\{g_n(y_1, \dots, y_{m+1}), n \geq 1\}$ is uniformly bounded, i.e., there exists $M > 0$ such that $|g_n(y_1, \dots, y_{m+1})| \leq M$, then when $c \geq 0$, we have

$$\limsup_{n \rightarrow \infty} \frac{1}{|T(n)|} \sum_{k=1}^n \sum_{t \in L_k} \{g_k(X_0^m(t)) - E_Q[g_k(X_0^m(t))|X_1^m(t)]\} \leq M(c+2\sqrt{c}), \quad P-a.e., \omega \in D(c). \quad (42)$$

Proof By (20) and (12) and the formula in line 2 of (22), we have

$$\begin{aligned} & \limsup_{n \rightarrow \infty} \frac{\lambda}{|T(n)|} \sum_{k=1}^n \sum_{t \in L_k} \{g_k(X_0^m(t)) - E_Q[g_k(X_0^m(t))|X_1^m(t)]\} \\ & \leq \limsup_{n \rightarrow \infty} \frac{1}{|T(n)|} \sum_{k=1}^n \sum_{t \in L_k} E_Q[e^{\lambda g_k(X_0^m(t))} - 1 - \lambda g_k(X_0^m(t))|X_1^m(t)] \\ & \quad + c \quad P-a.e., \omega \in D(c). \end{aligned} \quad (43)$$

By the hypothesis of the theorem and the inequality $e^x - 1 - x \leq |x|(e^{|x|} - 1)$, we have

$$e^{\lambda g_k(X_0^m(t))} - 1 - \lambda g_k(X_0^m(t)) \leq |\lambda| M (e^{|\lambda| M} - 1). \quad (44)$$

By (43) and (44)

$$\begin{aligned} & \limsup_{n \rightarrow \infty} \frac{\lambda}{|T(n)|} \sum_{k=1}^n \sum_{t \in L_k} \{g_k(X_0^m(t)) - E_Q[g_k(X_0^m(t))|X_1^m(t)]\} \\ & \leq |\lambda| M (e^{|\lambda| M} - 1) + c, \quad P-a.e., \omega \in D(c). \end{aligned} \quad (45)$$

When $\lambda > 0$, we have by (45)

$$\begin{aligned} & \limsup_{n \rightarrow \infty} \frac{1}{|T(n)|} \sum_{k=1}^n \sum_{t \in L_k} \{g_k(X_0^m(t)) - E_Q[g_k(X_0^m(t))|X_1^m(t)]\} \\ & \leq M(e^{\lambda M} - 1) + c/\lambda, \quad P-a.e., \omega \in D(c). \end{aligned} \quad (46)$$

Taking $\lambda = \frac{1}{M} \log(1 + \sqrt{c})$, and using the inequality

$$\log(1 + \sqrt{c}) \geq \frac{\sqrt{c}}{1 + \sqrt{c}}, \quad (47)$$

we have when $c > 0$

$$\begin{aligned} & \limsup_{n \rightarrow \infty} \frac{1}{|T^{(n)}|} \sum_{k=1}^n \sum_{t \in L_k} \{g_k(X_0^m(t)) - E_Q[g_k(X_0^m(t))|X_1^m(t)]\} \\ & \leq M\sqrt{c} + \frac{cM}{\log(1 + \sqrt{c})} \\ & \leq M(2\sqrt{c} + c), \quad P - a.e., \omega \in D(c). \end{aligned} \quad (48)$$

When $\lambda < 0$, it follows from (45) that

$$\begin{aligned} & \liminf_{n \rightarrow \infty} \frac{1}{|T^{(n)}|} \sum_{k=1}^n \sum_{t \in L_k} \{g_k(X_0^m(t)) - E_Q[g_k(X_0^m(t))|X_1^m(t)]\} \\ & \geq -M(e^{\lambda M} - 1) + c/\lambda, \quad P - a.e., \omega \in D(c). \end{aligned} \quad (49)$$

Taking $\lambda = -\frac{1}{M} \log(1 + \sqrt{c})$ in (49), and using (47), we have when $c > 0$

$$\begin{aligned} & \liminf_{n \rightarrow \infty} \frac{1}{|T^{(n)}|} \sum_{k=1}^n \sum_{t \in L_k} \{g_k(X_0^m(t)) - E_Q[g_k(X_0^m(t))|X_1^m(t)]\} \\ & \geq -M\sqrt{c} - \frac{cM}{\log(1 + \sqrt{c})} \\ & \geq -M(2\sqrt{c} + c), \quad P - a.e., \omega \in D(c). \end{aligned} \quad (50)$$

In a similar way, it can be shown that (48) and (50) also hold when $c = 0$. By (48) and (50), we have (42) holds. This completes the proof of this theorem. \square

Corollary 2 Under the conditions of Theorem 1, let $g(y_1, \dots, y_{m+1})$ be any function defined on G^{m+1} . Let $M = \max g(y_1, \dots, y_{m+1})$. Then when $c \geq 0$,

$$\limsup_{n \rightarrow \infty} \frac{1}{|T^{(n)}|} \left| \sum_{k=1}^n \sum_{t \in L_k} \{g(X_0^m(t)) - E_Q[g(X_0^m(t))|X_1^m(t)]\} \right| \leq M(c+2\sqrt{c}), \quad P - a.e., \omega \in D(c). \quad (51)$$

Proof Letting $g(y_1, \dots, y_{m+1}) = g_n(y_1, \dots, y_{m+1})$, $n \geq 1$ in Theorem 3, this corollary follows.

In the following, let $I_k(x) = \begin{cases} 1 & x = k \\ 0 & x \neq k \end{cases}$. Let $S_{T^{(n)} \setminus o'_m}(i_1, \dots, i_m)$ be the number of (i_1, \dots, i_m)

in the collection of $\{X_0^{m-1}(t), t \in T^{(n)} \setminus o'_m\}$, that is

$$S_{T^{(n)} \setminus o'_m}(i_1, \dots, i_m) = \sum_{k=0}^n \sum_{t \in L_k} I_{i_1}(X_{(m-1)_t}) \cdots I_{i_m}(X_t), \quad (52)$$

$S_{T^{(n)} \setminus o_m}(i_1, \dots, i_m, i_{m+1})$ be the number of $(i_1, \dots, i_m, i_{m+1})$ in the collection of $\{X_0^m(t), t \in T^{(n)} \setminus o_m\}$, that is

$$S_{T^{(n)} \setminus \mathcal{O}'_m}(i_1, \dots, i_m, i_{m+1}) = \sum_{k=1}^n \sum_{t \in L_k} I_{i_1}(X_{m_t}) \cdots I_{i_{m+1}}(X_t). \quad (53)$$

Corollary 3 Let $\{X_t, t \in T\}$ be defined as before. Then for all $i_1, \dots, i_{m+1} \in G$, $c \geq 0$, we have

$$\limsup_{n \rightarrow \infty} \left| \frac{S_{T^{(n)} \setminus \mathcal{O}'_m}(i_1, \dots, i_m)}{|T^{(n)}|} - \frac{1}{|T^{(n-1)}|} \sum_{l \in G} \sum_{k=0}^{n-1} \sum_{t \in L_k} I_l(X_{(m-1)_t}) \right. \\ \left. \cdot I_{i_1}(X_{(m-2)_t}) \cdots I_{i_{m-1}}(X_t) q_{k+1}(i_m | l, i_1, \dots, i_{m-1}) \right| \leq c + 2\sqrt{c}, \quad P - a.e., \omega \in D(c). \quad (54)$$

$$\limsup_{n \rightarrow \infty} \left| \frac{S_{T^{(n)} \setminus \mathcal{O}_m}(i_1, \dots, i_{m+1})}{|T^{(n)}|} - \frac{1}{|T^{(n-1)}|} \sum_{k=0}^{n-1} \sum_{t \in L_k} I_{i_1}(X_{(m-1)_t}) \right. \\ \left. \cdot I_{i_2}(X_{(m-2)_t}) \cdots I_{i_m}(X_t) q_{k+1}(i_{m+1} | i_1, \dots, i_m) \right| \leq c + 2\sqrt{c}, \quad P - a.e., \omega \in D(c). \quad (55)$$

Proof Letting $g(\gamma_1, \dots, \gamma_{m+1}) = I_{i_1}(\gamma_2) \cdots I_{i_m}(\gamma_{m+1})$ in Corollary 2.

$$\sum_{k=1}^n \sum_{t \in L_k} g(X_0^m(t)) = \sum_{k=1}^n \sum_{t \in L_k} I_{i_1}(X_{(m-1)_t}) \cdots I_{i_m}(X_t) \\ = S_{T^{(n)} \setminus \mathcal{O}'_m}(i_1, \dots, i_m) - I_{i_1}(X_{-(m-1)}) \cdots I_{i_m}(X_0), \quad (56)$$

and

$$\sum_{k=1}^n \sum_{t \in L_k} E_Q[g(X_0^m(t)) | X_1^m(t)] \\ = \sum_{k=1}^n \sum_{t \in L_k} \sum_{x_t \in G} g(X_1^m(t), x_t) q_k(x_t | X_1^m(t)) \\ = \sum_{k=1}^n \sum_{t \in L_k} \sum_{x_t \in G} I_{i_1}(X_{(m-1)_t}) \cdots I_{i_{m-1}}(X_{1_t}) I_{i_m}(x_t) q_k(x_t | X_1^m(t)) \\ = \sum_{k=1}^n \sum_{t \in L_k} I_{i_1}(X_{(m-1)_t}) \cdots I_{i_{m-1}}(X_{1_t}) q_k(i_m | X_1^m(t)) \\ = \sum_{l \in G} \sum_{k=1}^n \sum_{t \in L_k} I_l(X_{m_t}) I_{i_1}(X_{(m-1)_t}) \cdots I_{i_{m-1}}(X_{1_t}) q_k(i_m | l, i_1, \dots, i_{m-1}) \\ = N \sum_{l \in G} \sum_{k=0}^{n-1} \sum_{t \in L_k} I_l(X_{(m-1)_t}) I_{i_1}(X_{(m-2)_t}) \cdots I_{i_{m-1}}(X_t) q_{k+1}(i_m | l, i_1, \dots, i_{m-1}). \quad (57)$$

Noticing that $M = \max g(\gamma_1, \dots, \gamma_{m+1}) = 1$, $\lim_{n \rightarrow \infty} \frac{|T^{(n-1)}|}{|T^{(n)}|} = \frac{1}{N}$, by (56) and (57) and Corollary 2, (54) holds. Similarly, we let $g(\gamma_1, \dots, \gamma_{m+1}) = I_{i_1}(\gamma_1) \cdots I_{i_{m+1}}(\gamma_{m+1})$, (55) follows.

Corollary 4 Let $\{X_t, t \in T\}$ be defined as before.

$$\lim_{n \rightarrow \infty} \frac{1}{|T^{(n)}|} \sum_{k=1}^n \sum_{t \in L_k} \{g(X_0^m(t)) - E_Q[g(X_0^m(t)) | X_1^m(t)]\} = 0, \quad P - a.e., \omega \in D(0), \quad (58)$$

$$\lim_{n \rightarrow \infty} \left\{ \frac{S_{T^{(n)} \setminus \partial'_m}(i_1, \dots, i_m)}{|T^{(n)}|} - \frac{1}{|T^{(n-1)}|} \sum_{l \in G} \sum_{k=0}^{n-1} \sum_{t \in L_k} I_l(X_{(m-1)_t}) \right. \\ \left. \cdot I_{i_1}(X_{(m-2)_t}) \cdots I_{i_{m-1}}(X_t) q_{k+1}(i_m | l, i_1, \dots, i_{m-1}) \right\} = 0, \quad P - a.e., \omega \in D(0), \quad (59)$$

$$\lim_{n \rightarrow \infty} \left\{ \frac{S_{T^{(n)} \setminus \partial'_m}(i_1, \dots, i_{m+1})}{|T^{(n)}|} - \frac{1}{|T^{(n-1)}|} \sum_{k=0}^{n-1} \sum_{t \in L_k} I_{i_1}(X_{(m-1)_t}) \right. \\ \left. \cdot I_{i_2}(X_{(m-2)_t}) \cdots I_{i_m}(X_t) q_{k+1}(i_{m+1} | i_1, \dots, i_m) \right\} = 0, \quad P - a.e., \omega \in D(0). \quad (60)$$

If $P = Q$, then above equations hold $Q - a.e.$

Proof Letting $c = 0$ in Corollary 2 and Corollary 3, (58)-(60) follow from (51),(54) and (55). In particular, if $P = Q$, then $h(P|Q) = 0$, so (58)-(60) hold $P - a.e.$, hence hold $Q - a.e.$

Definition 3 Let $G = \{0, 1, \dots, b-1\}$ be a finite state space and

$$Q_1 = (q(j|i^m)), \quad j \in G, i^m \in G^m \quad (61)$$

be an m th-order transition matrix. Define a stochastic matrix as follows:

$$\bar{Q}_1 = (q(j^m|i^m)), \quad i^m, j^m \in G^m, \quad (62)$$

where

$$q(j^m|i^m) = \begin{cases} q(j_m|i^m), & \text{if } j_v = i_{v+1}, v = 1, 2, \dots, m-1, \\ 0, & \text{otherwise.} \end{cases} \quad (63)$$

Then \bar{Q}_1 is called an m -dimensional stochastic matrix determined by the m th-order transition matrix Q_1 .

Lemma 2 (see [16]). Let \bar{Q}_1 be an m -dimensional stochastic matrix determined by the m th-order transition matrix Q_1 . If the elements of Q_1 are all positive, that is

$$Q_1 = (q(j|i^m)), \quad q(j|i^m) > 0, \forall j \in G, i^m \in G^m, \quad (64)$$

then \bar{Q}_1 is ergodic.

Theorem 4 Let $\{X_t, t \in T\}$ be defined as Theorem 1. Let $S_{T^{(n)} \setminus \partial'_m}(i_1, \dots, i_m) = S_{T^{(n)} \setminus \partial'_m}(i^m)$, $S_{T^{(n)} \setminus \partial_m}(i_1, \dots, i_m, i_{m+1}) = S_{T^{(n)} \setminus \partial_m}(i^{m+1})$ and $f_n(\omega)$ defined by (52),(53) and (2), respectively. Let $h(P|Q)$ and $D(c)$ be defined by (4) and (12), respectively. Let the m th-order transition matrices defined by (6) be changeless with n , that is

$$Q_n = Q_1 = (q(j|i^m)), \quad (65)$$

or $\{X_t, t \in T\}$ is an m th-order homogeneous Markov chain indexed by tree T with the m th-order transition matrix Q_1 under the probability measure Q . Let the m -dimensional stochastic matrix \bar{Q}_1 determined by Q_1 be ergodic. Then for all $i_1, \dots, i_{m+1} \in G$, we have

$$\lim_{n \rightarrow \infty} \frac{S_{T^{(n)} \setminus \partial'_m}(i^m)}{|T^{(n)}|} = \pi(i^m), \quad P - a.e., \omega \in D(0). \quad (66)$$

$$\lim_{n \rightarrow \infty} \frac{S_{T(n) \setminus \{0_m\}}(i^{m+1})}{|T(n)|} = \pi(i^m)q(i_{m+1}|i^m), \quad P-a.e., \omega \in D(0). \quad (67)$$

$$\lim_{n \rightarrow \infty} f_n(\omega) = - \sum_{i^m \in G^m} \sum_{j \in G} \pi(i^m)q(j|i^m) \ln q(j|i^m), \quad P-a.e., \omega \in D(0). \quad (68)$$

where $\{\pi(i^m), i^m \in G^m\}$ is the stationary distribution determined by \bar{Q}_1 .

Proof Proof of Equation 66. Let $k^m = (k_1, \dots, k_m)$. If (65) holds, then we have by (63) and (52)

$$\begin{aligned} & \sum_{l \in G} \sum_{k=0}^{n-1} \sum_{t \in L_k} I_l(X_{(m-1)_t}) I_{i_1}(X_{(m-2)_t}) \cdots I_{i_{m-1}}(X_t) q_{k+1}(i_m|l, i_1, \dots, i_{m-1}) \\ &= \sum_{l \in G} \sum_{k=0}^{n-1} \sum_{t \in L_k} I_l(X_{(m-1)_t}) I_{i_1}(X_{(m-2)_t}) \cdots I_{i_{m-1}}(X_t) q(i_m|l, i_1, \dots, i_{m-1}) \\ &= \sum_{l \in G} S_{T(n-1) \setminus \{0'_m\}}(l, i_1, \dots, i_{m-1}) q(i_m|l, i_1, \dots, i_{m-1}) \\ &= \sum_{k^m \in G^m} S_{T(n-1) \setminus \{0'_m\}}(k^m) q(i^m|k^m). \end{aligned} \quad (69)$$

By (59) and (69), we have

$$\lim_{n \rightarrow \infty} \left\{ \frac{S_{T(n) \setminus \{0'_m\}}(i^m)}{|T(n)|} - \frac{1}{|T(n-1)|} \sum_{k^m \in G^m} S_{T(n-1) \setminus \{0'_m\}}(k^m) q(i^m|k^m) \right\} = 0, \quad P-a.e., \omega \in D(0). \quad (70)$$

Multiplying (70) by $q(j^m|i^m)$, adding them together for $i^m \in G^m$, and using (70) once again, we have

$$\begin{aligned} 0 &= \sum_{i^m \in G^m} q(j^m|i^m) \cdot \lim_{n \rightarrow \infty} \left\{ \frac{S_{T(n) \setminus \{0'_m\}}(i^m)}{|T(n)|} - \frac{1}{|T(n-1)|} \sum_{k^m \in G^m} S_{T(n-1) \setminus \{0'_m\}}(k^m) q(i^m|k^m) \right\} \\ &= \lim_{n \rightarrow \infty} \left\{ \sum_{i^m \in G^m} \frac{S_{T(n) \setminus \{0'_m\}}(i^m)}{|T(n)|} q(j^m|i^m) - \frac{S_{T(n+1) \setminus \{0'_m\}}(j^m)}{|T(n+1)|} \right\} \\ &\quad + \lim_{n \rightarrow \infty} \left\{ \frac{S_{T(n+1) \setminus \{0'_m\}}(j^m)}{|T(n+1)|} - \frac{1}{|T(n-1)|} \sum_{k^m \in G^m} S_{T(n-1) \setminus \{0'_m\}}(k^m) \sum_{i^m \in G^m} q(j^m|i^m) q(i^m|k^m) \right\} \\ &= \lim_{n \rightarrow \infty} \left\{ \frac{S_{T(n+1) \setminus \{0'_m\}}(j^m)}{|T(n+1)|} - \frac{1}{|T(n-1)|} \sum_{k^m \in G^m} S_{T(n-1) \setminus \{0'_m\}}(k^m) q^{(2)}(j^m|k^m) \right\}, \quad P-a.e., \omega \in D(0). \end{aligned}$$

By induction, we have

$$\lim_{n \rightarrow \infty} \left\{ \frac{S_{T(n+N) \setminus \{0'_m\}}(j^m)}{|T(n+N)|} - \frac{1}{|T(n-1)|} \sum_{k^m \in G^m} S_{T(n-1) \setminus \{0'_m\}}(k^m) q^{(N+1)}(j^m|k^m) \right\} = 0, \quad P-a.e., \omega \in D(0). \quad (71)$$

where $q^{(h)}(j^m|k^m)$ is the h th step probability determined by \bar{Q}_1 . We have by ergodicity

$$\lim_{N \rightarrow \infty} q^{(N+1)}(j^m|k^m) = \pi(j^m), \quad \forall k^m \in G^m, \quad (72)$$

and $\sum_{k^m \in G^m} S_{T(n-1) \setminus \{0'_m\}}(k^m) = |T(n-1)| - (m-1)$. (66) follows from (71) and (72). By (66) and (60), Equation 67 follows easily.

Proof of Equation 68. By (66) and (53), we have

$$\begin{aligned}
 & \sum_{k=1}^n \sum_{t \in L_k} H[q_k(0|X_1^m(t)), \dots, q_k(b-1|X_1^m(t))] \\
 &= \sum_{k=1}^n \sum_{t \in L_k} H[q(0|X_1^m(t)), \dots, q(b-1|X_1^m(t))] \\
 &= - \sum_{k=1}^n \sum_{t \in L_k} \sum_{j \in G} q(j|X_1^m(t)) \ln q(j|i^m) \\
 &= - \sum_{k=1}^n \sum_{t \in L_k} \sum_{j \in G} \sum_{i^m \in G^m} I_{i_1}(X_{m_t}) \cdots I_{i_m}(X_{1_t}) q(j|i^m) \ln q(j|i^m) \\
 &= -N \sum_{k=0}^{n-1} \sum_{t \in L_k} \sum_{j \in G} \sum_{i^m \in G^m} I_{i_1}(X_{(m-1)_t}) \cdots I_{i_m}(X_t) q(j|i^m) \ln q(j|i^m) \\
 &= -N \sum_{j \in G} \sum_{i^m \in G^m} S_{T^{(n-1)} \setminus o'_m}(i^m) q(j|i^m) \ln q(j|i^m).
 \end{aligned} \tag{73}$$

Noticing that $\lim_{n \rightarrow \infty} \frac{|T^{(n-1)}|}{|T^{(n)}|} = \frac{1}{N}$, by (39), (73) and (66), Equation 68 follows. \square

3. Shannon-McMillan Theorem

Theorem 5 Let $\{X_t, t \in T\}$ be a G -valued m th-order nonhomogeneous Markov chain indexed by an m rooted Cayley tree under the probability measure Q with initial distribution (5) and m th-order transition matrices (6). Let $S_{T^{(n)} \setminus o'_m}(i^m)$, $S_{T^{(n)} \setminus o_m}(i^{m+1})$ and $f_n(\omega)$ be defined as before. Let

$$Q_n = Q_1 = (q(j|i^m)), \quad q(j|i^m) > 0, \quad \forall i^m \in G^m, j \in G, \tag{74}$$

be another positive m th-order transition matrix. Let \bar{Q}_1 be an m dimension transition matrix determined by Q_1 . If

$$\lim_{n \rightarrow \infty} q_n(j|i^m) = q(j|i^m), \quad \forall i^m \in G^m, j \in G, \tag{75}$$

then

$$\lim_{n \rightarrow \infty} \frac{S_{T^{(n)} \setminus o'_m}(i^m)}{|T^{(n)}|} = \pi(i^m), \quad P - a.e. \quad \omega \in D(0), \tag{76}$$

$$\lim_{n \rightarrow \infty} \frac{S_{T^{(n)} \setminus o_m}(i^{m+1})}{|T^{(n)}|} = \pi(i^m) q(i_{m+1}|i^m), \quad P - a.e. \quad \omega \in D(0), \tag{77}$$

$$\lim_{n \rightarrow \infty} f_n(\omega) = - \sum_{i^m \in G^m} \sum_{j \in G} q(j|i^m) \ln q(j|i^m), \quad P - a.e. \quad \omega \in D(0), \tag{78}$$

where $\{\pi(i^m), i^m \in G^m\}$ is the stationary distribution determined by \bar{Q}_1 . In particular, if $P = Q$, then above equations hold $Q - a.e.$

Proof By (59), (75), (52) and (66), (76) follows immediately. Similarly, by (60), (75), and (53), (77) follows. It follows from (75) and Cesaro average that

$$\lim_{n \rightarrow \infty} \frac{1}{|T^{(n)}|} \sum_{k=1}^n \sum_{t \in L_k} |q_k(j|i^m) \ln q_k(j|i^m) - q(j|i^m) \ln q(j|i^m)| = 0, \quad \forall i^m \in G^m, j \in G. \tag{79}$$

Notice that

$$\begin{aligned} & \left| \frac{1}{|T^{(n)}|} \sum_{k=1}^n \sum_{t \in L_k} H[q_k(0|X_1^m(t)), \dots, q_k(0|X_1^m(t))] - \frac{1}{|T^{(n)}|} \sum_{k=1}^n \sum_{t \in L_k} H[q(0|X_1^m(t)), \dots, q(0|X_1^m(t))] \right| \\ &= \left| -\frac{1}{|T^{(n)}|} \sum_{k=1}^n \sum_{t \in L_k} \sum_{j \in G} q_k(j|X_1^m(t)) \ln q_k(j|X_1^m(t)) + \frac{1}{|T^{(n)}|} \sum_{k=1}^n \sum_{t \in L_k} \sum_{j \in G} q(j|X_1^m(t)) \ln q(j|X_1^m(t)) \right| \quad (80) \\ &\leq \frac{1}{|T^{(n)}|} \sum_{k=1}^n \sum_{t \in L_k} \sum_{j \in G} \sum_{i^m \in C^m} |q_k(j|i^m) \ln q_k(j|i^m) - q(j|i^m) \ln q(j|i^m)|. \end{aligned}$$

By (79), (80), (39), (73) and (66), (78) follows. In particular, if $P = Q$, then $h(P|Q) = 0$. (76), (77) and (78) also holds $P - a.e.$ \square

Acknowledgements

The authors would like to thank the referees for their valuable suggestions and comments. This work is supported by the Research Foundation for Advanced Talents of Jiangsu University (11JDG116) and the National Natural Science Foundation of China (11071104, 11171135, 71073072).

Author details

¹Faculty of Science, Jiangsu University, Zhenjiang 212013, China ²Department of Mathematics, Chaohu University, Chaohu 238000, China

Authors' contributions

ZS, WY and LT carried out the design of the study and performed the analysis. WP participated in its design and coordination. All authors read and approved the final manuscript.

Competing interests

The authors declare that they have no competing interests.

Received: 21 March 2011 Accepted: 4 January 2012 Published: 4 January 2012

References

- Liu, W, Wang, LY: The Markov approximation of the random field on Cayley tree and a class of small deviation theorems. *Stat Probab Lett.* **63**, 113–121 (2003). doi:10.1016/S0167-7152(03)00058-0
- Shi, ZY, Yang, WG: Some limit properties for the m th-order nonhomogeneous Markov chains indexed by an m rooted Cayley tree. *Stat Probab Lett.* **80**, 1223–1233 (2010). doi:10.1016/j.spl.2010.03.020
- Benjamini, I, Peres, Y: Markov chains indexed by trees. *Ann Probab.* **22**, 219–243 (1994). doi:10.1214/aop/1176988857
- Berger, T, Ye, Z: Entropic aspects of random fields on trees. *IEEE Trans Inf Theory.* **36**, 1006–1018 (1990). doi:10.1109/18.57200
- Pemantle, R: Andomorphism invariant measure on tree. *Ann Probab.* **20**, 1549–1566 (1992). doi:10.1214/aop/1176989706
- Ye, Z, Berger, T: Ergodic, regularity and asymptotic equipartition property of random fields on trees. *Combin Inf Syst Sci.* **21**, 157–184 (1996)
- Ye, Z, Berger, T: Information measure for discrete random field. Science, Beijing. (1998)
- Yang, WG, Liu, W: Strong law of large numbers for Markov chains fields on a bethe tree. *Stat Probab Lett.* **49**, 245–250 (2000). doi:10.1016/S0167-7152(00)00053-5
- Yang, WG: Some limit properties for Markov chains indexed by a homogeneous tree. *Stat Probab Lett.* **65**, 241–250 (2003). doi:10.1016/j.spl.2003.04.001
- Yang, WG, Ye, Z: The asymptotic equipartition property for Markov chains indexed by a Homogeneous tree. *IEEE Trans Inf Theory.* **53**(9), 3275–3280 (2007)
- Huang, HL, Yang, WG: Strong law of large numbers for Markov chains indexed by an infinite tree with uniformly bounded degree. *Sci China Ser A Math.* **51**(2), 195–202 (2008). doi:10.1007/s11425-008-0015-1
- Shi, ZY, Yang, WG: Some limit properties of random transition probability for second-order nonhomogeneous Markov chains indexed by a tree. *J Inequal Appl* ID 503203. (2009)
- Peng, WC, Yang, WG, Wang, B: A class of small deviation theorems for functional of random fields on a homogeneous tree. *J Math Anal Appl.* **361**, 293–301 (2009)
- Yang, WG: A class of small deviation theorems for the sequences of N -valued random variables with respect to m th-order nonhomogeneous Markov chains. *Acta Mathematica Scientia* **29A**(2), 517–527 (2009). in Chinese
- Gray, RM: Entropy and Information Theory. Springer, New York (1990)
- Yang, WG, Liu, W: The asymptotic equipartition property for M th-order nonhomogeneous Markov information sources. *IEEE Trans Inf Theory.* **50**(12), 3326–3330 (2004). doi:10.1109/TIT.2004.838339

doi:10.1186/1029-242X-2012-1

Cite this article as: Shi et al.: A class of small deviation theorems for the random fields on an m rooted Cayley tree. *Journal of Inequalities and Applications* 2012 **2012**:1.